

International Journal of Population Data Science

Journal Website: www.ijpds.org



Implementing Data Governance with Multi-Modal Privacy-Preserving Record Linkages between Restricted and Public Open Enclaves

Kenneth Gersing¹, Sam Michael¹, Fred Prior², Joel Saltz³, Richard Moffitt⁴, Sara Rogovin⁵, Jasmin Phua⁵, Shaun Grannis⁶, Ben Amor⁷, Andrew Girvin⁶, and Ahmad Baghal²

¹National Institutes of Health

²University of Arkansas Medical Sciences

³Stony Brook University

⁴Emory University

⁵Datavant

⁶Regenstrief Institute

⁷Palantir Technologies

Data centralization is increasingly non-viable in the changing landscape of data privacy and data governance. New models of data sharing are needed, in particular when privacy-preserving record linkages (PPRL) are being used to connect data within restricted data enclaves and public open enclaves.

We share a privacy-preserving data-sharing infrastructure that employs a linkage honest broker implementation of PPRL connecting structured electronic health record data between the National Center for Advancing Translational Science (NCATS) National Clinical Cohort Collaborative (N3C), a restricted data enclave, with the National Cancer Institute (NCI) Cancer Imaging Archive (TCIA), a public open enclave.

In particular we will address:

- Data governance models needed when linking data from restricted vs. public enclaves
- Specific privacy-preserving consideration when linking to imaging data which contain specific structured data and imaging considerations
- Privacy-preserving data aggregation when combining multi-modal data from restricted data enclaves with publicly discoverable imaging data

